Network Working Group
Request for Comments: 985

Network Technical Advisory Group
NSF
May 1986

Requirements for Internet Gateways -- Draft

Status of this Memo

This RFC summarizes the requirements for gateways to be used on networks supporting the DARPA Internet protocols. While it applies specifically to National Science Foundation research programs, the requirements are stated in a general context and are believed applicable throughout the Internet community. This document was prepared by the Gateway Requirements Subcommittee of the NSF Network Technical Advisory Group in cooperation with the Internet Activities Board, Internet Architecture Task Force and Internet Engineering Task Force. It requests discussion and suggestions for improvements. Distribution of this memo is unlimited.

The purpose of this document is to present guidance for vendors offering products that might be used or adapted for use in an Internet application. It enumerates the protocols required and gives references to RFCs and other documents describing the current specifications. In a number of cases the specifications are evolving and may contain ambiguous or incomplete information. In these cases further discussion giving specific guidance is included in this document. Specific policy issues relevant to the NSF scientific networking community are summarized in an Appendix.

****

This is a DRAFT edition of this statement of gateway requirements. Comments are sought on this document for consideration and possibly incorporated in the final edition. Comments are especially sought from those actually developing gateways, particular vendors and potential vendors of gateways. The period for comments is 90 days ending 15-Aug-86, at which time revised edition will be issued with a new RFC number.

****

Suggestions and comments on this document can be sent to the subcommittee chairman Dave Mills (mills@usc-isid.arpa), or NTAG committee chairman Dave Farber (farber@huey.udel.edu). The subcommittee members, present affiliations and Internet mailboxes are as follows:

Hank Dardy, NRL	dardy@nrl.arpa
Dave Farber, U Delaware	farber@huey.udel.edu
Dennis Jennings, JVNC	jennings%pucc.bitnet@wiscvm.wisc.edu

Larry Landweber, U Wisconsin landweber@rsch.wisc.edu
Tony Lauck, DEC rhea!bergil!lauck@decwrl.arpa
Dave Mills (Chairman), Linkabit mills@usc-isid.arpa
Dennis Perry, DARPA/IPTO perry@ipto.arpa

The subcommittee wishes to thank the following additional contributors and invited referees:

Len Bosack, Stanford U/CISCO bosack@su-score.arpa
Bob Braden, ISI braden@isi-braden.arpa
Hans-Werner Braun, U Michigan hwb@gw.umich.edu
Noel Chiappa, MIT/Proteon jnc@proteon.arpa
Doug Comer, Purdue U dec@cs.purdue.edu
Ira Fuchs, Princeton U fuchs%pucc.bitnet@wiscvm.wisc.edu
Ed Krol, U Illinois krol%uiucvmd.bitnet@wiscvm.wisc.edu
Barry Leiner, RIACS leiner@riacs.arpa
Mike Muuss, BRL mike@brl.arpa
Ron Natalie, BRL ron@brl.arpa
Harvey Newman, CIT newman@cit-hex.arpa
Jon Postel, ISI postel@usc-isib.arpa
Marshall Rose, NRTC mrose@nrtc-gremlin.northrop.com
Jeff Schiller, MIT jis@bitsy.mit.edu
Lixia Zhang, MIT lixia@xx.lcs.mit.edu

1. Introduction

The following sections are intended as an introduction and background for those unfamiliar with the DARPA Internet architecture and the Internet gateway model. General background and discussion on the Internet architecture and supporting protocol suite can be found in the DDN Protocol Handbook [25] and ARPANET Information Brochure [26], both available from the Network Information Center, SRI International, Menlo Park, CA 94025. Readers familiar with these concepts can proceed directly to Section 2.

1.1. The DARPA Internet Architecture

The DARPA Internet system consists of a number of gateways and networks that collectively provide packet transport for hosts subscribing to the DARPA Internet protocol architecture. These protocols include the Internet Protocol (IP), Internet Control Message Protocol (ICMP), Transmission Control Protocol (TCP) and application protocols depending upon them. All protocols use IP as the basic packet-transport mechanism. IP is a datagram, or connectionless, service and includes provision for service specification, fragmentation/reassembly and security information. ICMP is considered an integral part of IP, although it is

architecturally layered upon it. ICMP provides error reporting, flow control and first-hop gateway redirection. Reliable data delivery is provided in the protocol suite by TCP, which provides end-end retransmission, resequencing and connection control. Connectionless service is provided by the User Datagram Protocol (UDP).

The Internet community presently includes several thousand hosts connected to over 400 networks with about 120 gateways. There are now well over 2400 hosts registered in the ARPA domain alone and an unknown number registered in other domains, with the total increasing at about ten percent each month. Many of the hosts, gateways and networks in the Internet community are administered by civil organizations, including universities, research laboratories and equipment manufacturers. Most of the remainder are administered by the US DoD and considered part of the DDN Internet, which presently consists of three sets of networks: the experimental segment, or ARPANET, the unclassified segment, or MILNET, and the classified segment, which does not yet have a collective name.

The Internet model includes constituent networks, called local networks to distinguish them from the Internet system as a whole, which are required only to provide datagram (connectionless) transport. This requires only best-effort delivery of individual packets, or datagrams. Each datagram carries 32-bit source and destination addresses, which are encoded in three formats providing a two-part address, one of which is the local-network number and the other the host number on that local net. According to the Internet service specification, datagrams can be delivered out of order, be lost or duplicated and/or contain errors. In those networks providing connection-oriented service the extra reliability provided by virtual circuits enhances the end-end robustness of the system, but is not strictly necessary.

Local networks are connected together in the Internet model by means of Internet gateways. These gateways provide datagram transport only and normally seek to minimize the state information necessary to sustain this service in the interest of routing flexibility and robustness. In the conventional model the gateway has a physical interface and address on each of the local nets between which it provides forwarding services. The gateway also participates in one or more distributed routing or reachability algorithm such as the Gateway-Gateway Protocol (GGP) or Exterior Gateway Protocol (EGP) in order to maintain its routing tables.

1.2. The Internet Gateway Model

An Internet gateway is a self-contained, stand-alone packet switch that performs the following functions:

1. Interfaces to two or more packet-switching networks, including encapsulation, address transformation and flow control.

2. Conforms to specific DARPA Internet protocols specified in this document, including the Internet Protocol (IP), Internet Control Message Protocol (ICMP), Exterior Gateway Protocol (EGP) and others as necessary.

3. Supports an interior gateway protocol (IGP) reachability or routing algorithm in cases of multiple gateways operating as a system. Supports the EGP reachability algorithm to exchange routes between systems, in particular the DARPA "core" system operated by BBN.

4. Receives and forwards Internet datagrams consistent with good engineering practice in the management of resources, congestion control and fairness. Recognizes various error conditions and generates ICMP error and information messages as required.

5. Provides system support facilities, including loading, debugging, status reporting, exception reporting and control.

In some configurations gateways may be connected to packet-switching local nets that provide generic local-net routing, error-control and resource-management functions. In others gateways may be directly connected via serial lines, so that these functions must be provided by the gateways themselves.

There are three typical scenarios that should be addressed by gateway vendors:

1. National or regional network. Gateways of this class should be capable of switching multiple continuous flows in the 1.5-Mbps range at rates to several thousand packets per second. They will be high-performance, possibly redundant, multiple-processor devices, probably procured as a system and operated remotely from a regional or national monitoring center. The design of these gateways should emphasize high aggregate throughput, throughput-sensitive

resource management and very high reliability. The typical application would be an NSF backbone net or one of the consortium or regional nets.

2. Campus network. Gateways of this class should be capable of switching some burst flows at 10-Mbps (Ethernets, etc.), together with some flows in the 64-Kbps range or lower, at rates to perhaps several thousand packets per second. They will be medium-performance devices, probably competitively procured from different vendors for each campus and operated from a campus computing center. The design of these gateways should emphasize low average delay and good burst performance, together with delay and type-of-service sensitive resource management. Their chief function might be to interconnect various LANs and campus computing resources, including a high-speed interconnect to a national or regional net. An important factor will be a very flexible routing mechanism, since these gateways may have to select among several backbone nets based on cost/performance considerations.

3. Department network. Gateways of this class should be capable of switching a small number of burst flows at 10-Mbps (Ethernets, etc.), together with a small number of flows in the range 64-Kbps or lower, at rates of a few hundred packets per second. They will be medium-performance devices procured from a variety of vendors and used for protocol-matching, LAN repeaters and as general utility packet switches. They will probably be locally maintained by the various users and not be used as transit switches.

It is important to realize that Internet gateways normally operate in an unattended mode, but that equipment and software faults can affect the entire Internet. While some of the above scenarios involve positive control of some gateways from a monitoring center, usually via a path involving other networks and Internet gateways, others may involve much less formal control procedures. Thus the gateways must be highly robust and be expected to operate, possibly in a degraded state, under conditions of extreme congestion or failure of network resources.

2. Protocols Required

The Internet architecture uses datagram gateways to interconnect networks and subnetworks. These gateways function as intermediate systems (IS) with respect to the ISO connectionless network model and incorporate defined packet formats, routing algorithms and related procedures. In the following it is assumed the protocol implementation supports the full protocol, including all required options, with exceptions only as noted.

2.1. Internet Protocol (IP)

This is the basic datagram protocol used in the Internet system. It is described in RFC-791 [1] and also MIL-STD-1777 [5], both of which are intended to describe the same standard, but in quite different words.

With respect to current gateway requirements the following can be ignored, although they may be required in future: Type of Service field, Security option, Stream ID option and Timestamp option. However, if recognized, the interpretation of these quantities must conform to the standard specification.

Note that the Internet gateway model does not require that the gateway reassemble IP datagrams with destination address other than the gateway itself. However, in the case of those protocols in which the gateway directly participates as a peer, including routing and monitor/control protocols, the gateway may have to reassemble datagrams addressed to it. This consideration is most pertinent to EGP.

Note that, of the five classes of IP addresses. Class-A through Class-E, Class-D and Class-E addresses are reserved for experimental use. A gateway which is not participating in these experiments should ignore all packets with a Class-D or Class-E destination IP address. No ICMP Destination Unreachable or ICMP Redirect messages should result from receiving such packets.

2.2. Internet Control Message Protocol (ICMP)

This is an auxiliary protocol used to convey advice and error messages and is described in RFC-792 [2].

The distinction between subnets of a subnetted network, which depends on an arbitrary mask as described in RFC-950 [21], is in general not visible outside that network. This distinction is important in the case of certain ICMP messages, including the ICMP

Destination Unreachable and ICMP Redirect messages. The ICMP Destination Unreachable message is sent by a gateway in response to a datagram which cannot be forwarded because the destination is unreachable or down. A choice of several types of these messages is available, including one designating the destination network and another the destination host. However, the span of addresses implied by the former is ill-defined unless the subnet mask is known to the sender, which is in general not the case. It is recommended that use of the ICMP Destination Network Unreachable messages be avoided. Instead, an ICMP Destination Host Unreachable message should be sent for each distinct unreachable IP address.

The ICMP Redirect message is sent by a gateway to a host in order to change the address used by the host for a designated host or net. A choice of four types of messages is available, depending on whether it applies to a particular host, network or service. As in the previous case, these distinctions may depend upon the subnet mask. As in the above case, it is recommended that the use of ICMP messages implying a span of addresses (e.g. net unreachable, net redirect) be avoided in favor of those implying specific addresses (e.g. host unreachable, host redirect).

The ICMP Source Quench message has been the subject of much controversy. It is not considered realistic at this time to specify in detail the conditions under which this message is to be generated or interpreted by a host or gateway.

New host and gateway implementations are expected to support the ICMP Address Mask messages described in RFC-950. It is highly desirable, although not required, to provide correct data for ICMP Timestamp messages, which have been found useful in network debugging and maintenance.

2.3. Exterior Gateway Protocol (EGP)

This is the basic protocol used to exchange information between gateway systems of the Internet and is described in RFC-904 [11]. However, EGP as presently specified is an asymmetric protocol with only the "non-core" procedures defined in RFC-904. There are at present no "core" procedures specified, which would be necessary for a stand-alone Internet. RFC-975 [27] suggests certain modifications leading to a symmetric model; however, this is not an official specification.

In principle, a stand-alone Internet can be built with non-core EGP gateways using the EGP distance field to convey some metric

such as hop count. However, the use of EGP in this way as a routing algorithm is discouraged, since typical implementations adapt very slowly to changing topology and have no loop-protection features.

The EGP model requires each gateway belong to an autonomous system of gateways. If a routing algorithm is operated in one or more gateways of an autonomous system, its data base must be coupled to the EGP implementation in such a way that, when a net is declared down by the routing algorithm, the net is also declared down via EGP to other autonomous systems. This requirement is designed to minimize spurious traffic to "black holes" and insure fair utilization of the resources on other systems.

There are no peer-discovery or authentication procedures defined in the present EGP specification and no defined interpretation of the distance fields in the update messages, although such procedures may be defined in future (see RFC-975). There is currently no guidance on the selection of polling parameters and no specific recovery procedures in case of certain error messages (e.g. "administratively prohibited"). It is recommended that EGP implementations include provisions to initialize these parameters as part of the monitoring and control procedures and that changing these procedures not require recompilation or rebooting the gateway.

2.4. Address Resolution Protocol (ARP)

This is an auxiliary protocol used to manage the address-translation function between hardware addresses in a local-net environment and Internet addresses and described in RFC-826 [4]. However, there are a number of unresolved issues having to do with subnets and response to addresses not in the same subnet or net. These issues, which are intertwined with ICMP and various gateway models, are discussed in Appendix A.

3. Subnets

The concept of subnets was introduced in order to allow arbitrary complexity of interconnected LAN structures within an organization, while insulating the Internet system against explosive growth in network numbers and routing complexity. The subnet architecture, described in RFC-950 [21], is intended to specify a standard approach that does not require reconfiguration for host implementations, regardless of subnetting scheme. The document also specifies a new

ICMP Address Mask message, which a gateway can use to specify certain details of the subnetting scheme to hosts and is required in new host and gateway implementations.

The current subnet specification RFC-950 does not describe the specific procedures to be used by the gateway, except by implication. It is recommended that a (sub)net address and address mask be provided for each network interface and that these values be established as part of the gateway configuration procedure. It is not usually necessary to change these values during operation of any particular gateway; however, it should be possible to add new gateways and/or (sub)nets and make other configuration changes to a gateway without taking the entire network down.

4. Local Network Interface

The packet format used for transmission of datagrams on the various subnetworks is described in a number of documents summarized below.

4.1. Public data networks via X.25

The formats specified for public data networks via X.25 access are described in RFC-877 [8]. Datagrams are transmitted over standard level-3 virtual circuits as complete packet sequences. Virtual circuits are usually established dynamically as required and time out after a period of no traffic. Retransmission, resequencing and flow control are performed by the network for each virtual circuit and by the LAPB link-level protocol. Multiple parallel virtual circuits are often used in order to improve the utilization of the subscriber access line, which can result in random resequencing. The correspondence between Internet and X.121 addresses is usually established by table-lookup. It is expected that this will be replaced by some sort of directory procedure in future.

4.2. ARPANET via 1822 Local Host, Distant Host or HDLC Distant Host

The formats specified for ARPANET networks via 1822 access are described in BBN Report 1822 [3], which includes the procedures for several subscriber access methods. The Local Host (LH) and Very Distant Host (VDH) methods are not recommended for new implementations. The Distant Host (DH) method is used when the host and IMP are separated by not more than about 2000 feet of cable, while the HDLC Distant Host is used for greater distances where a modem is required. Retransmission, resequencing and flow control are performed by the network and by the HDLC link-level protocol, when used. While the ARPANET 1822 protocols are widely

used at present, they are expected to be eventually overtaken by the DDN Standard X.25 protocol (see below) and the new PSN End-to-End Protocol described in RFC-979 [29].

While the cited report gives details of the various ARPANET subscriber access methods, it specifies neither the IP packet encapsulation format nor address mappings. While these are generally straightforward and easy to implement, the details involve considerations beyond the scope of readily accessable documentation. Potential vendors are encouraged to contact one of the individuals listed at the beginning of this document for further information.

Gateways connected to ARPANET/MILNET IMPs must incorporate features to avoid host-port blocking (RFNM counting) and to detect and report (as ICMP Unreachable messages) the failure of destination hosts or gateways.

4.3. ARPANET via DDN Standard X.25

The formats specified for ARPANET networks via X.25 are described in the Defense Data Network X.25 Host Interface Specification [6]. This document describes two sets of procedures, the DDN Basic X.25 and the DDN Standard X.25, but only the latter is suitable for use in the Internet system. The DDN Standard X.25 procedures are similar to the public data subnetwork X.25 procedures, except in the address mappings. Retransmission, resequencing and flow control are performed by the network and by the LAPB link-level protocol.

4.4. Ethernets

The formats specified for Ethernet networks are described in RFC-894 [10]. Datagrams are encapsulated as Ethernet packets with 48-bit source and destination address fields and a 16-bit type field. Address translation between Ethernet addresses and Internet addresses is managed by the Address Resolution Protocol, which is required in all Ethernet implementations. There is no explicit retransmission, resequencing or flow control. although most hardware interfaces will retransmit automatically in case of collisions on the cable.

It is expected that amendments will be made to this specification as the result of IEEE 802.3 evolution. See RFC-948 [20] for further discussion and recommendations in this area. Note also that the IP broadcast address, which has primary application to Ethernets and similar technologies that support an inherent

broadcast function, has an all-ones value in the host field of the IP address. Some early implementations chose the all-zeros value for this purpose, which is presently not in conformance with the definitive specification RFC-950 [21].

See Appendix A for further considerations.

4.5. Serial-Line Protocols

Gateways may be used as packet switches in order to build networks. In some configurations gateways may be interconnected with each other and some hosts by means of serial asynchronous or synchronous lines, with or without modems. When justified by the expected error rate and other factors, a link-level protocol may be required on the serial line. While there is no requirement that a particular standard protocol be used for this, it is recommended that standard hardware and protocols be used, unless a convincing reason to the contrary exists. In order to support the greatest variety of configurations, it is recommended that some variation on full X.25 (i.e. "symmetric mode") be used where resources permit; however, X.25 LAPB would also be acceptable where requirements permit. In the case of asynchronous lines no clear choice is apparent.

5. Interoperability

In order to assure interoperability between gateways procured from different vendors, it is necessary to specify points of protocol demarcation. With respect to interoperability of the routing function, this is specified as EGP. All gateway systems must include one or more gateways which support EGP with a core gateway, as described in RFC-904 [11]. It is desirable that these gateways be able to operate in a mode that does not require a core gateway or system. Additional discussion on these issues can be found in RFC-975 [27].

With respect to the interoperability at the network layer and below, two points of protocol demarcation are specified, one for Ethernets and the other for serial lines. In the case of Ethernets the protocols are as specified in Section 4.4 and Appendix A of this document. For serial lines between gateways of different vendors, the protocols are specified in Section 4.5 of this document. Exceptions to these requirements may be appropriate in some cases.

6. Subnetwork Architecture

It is recognized that gateways may also function as general packet switches to build networks of modest size. This requires additional functionality in order to manage network routing, control and configuration. While it is beyond the scope of this document to specify the details of the mechanisms used in any particular, perhaps proprietary, architecture, there are a number of basic requirements which must be provided by any acceptable architecture.

6.1. Reachability Procedures

The architecture must provide a robust mechanism to establish the operational status of each link and node in the network, including the gateways, the links connecting them and, where appropriate, the hosts as well. Ordinarily, this requires at least a link-level reachability protocol involving a periodic exchange of hello messages across each link. This function might be intrinsic to the link-level protocols used (e.g. LAPB, DDCMP). However, it is in general ill-advised to assume a host or gateway is operating correctly if its link-level reachability protocol is operating correctly. Additional confirmation is required in the form of an operating routing algorithm or peer-level reachability protocol, such as used in EGP.

Failure and restoration of a link and/or gateway are considered network events and must be reported to the control center. It is desirable, although not required, that reporting paths not require correct functioning of the routing algorithm itself.

6.2. Routing Algorithm

It has been the repeated experience of the Internet community participants that the routing mechanism, whether static or dynamic, is the single most important engineering issue in network design. In all but trivial network topologies it is necessary that some degree of routing dynamics is vital to successful operation, whether it be affected by manual or automatic means or some combination of both. In particular, if routing changes are made manually, the changes must be possible without taking down the gateways for reconfiguration and, preferably, be possible from a remote site such as a control center.

It is not likely that all nets can be maintained from a full-service control center, so that automatic-fallback or rerouting features may be required. This must be considered the normal case, so that systems of gateways operating as the only

packet switches in a network would normally be expected to have a routing algorithm with the capability of reacting to link and other gateway failures and changing the routing automatically. Following is a list of features considered necessary:

1. The algorithm must sense the failure or restoration of a link or other gateway and switch to appropriate paths within an interval less than the typical TCP user timeout (one minute is a safe assumption).

2. The algorithm must never form routing loops between neighbor gateways and must contain provisions to avoid and suppress routing loops that may form between non-neighbor gateways. In no case should a loop persist for longer than an interval greater than the typical TCP user timeout.

3. The control traffic necessary to operate the routing algorithm must not significantly degrade or disrupt normal network operation. Changes in state which might momentarily disrupt normal operation in a local area must not cause disruption in remote areas of the network.

4. As the size of the network increases, the demand on resources must be controlled in an efficient way. Table lookups should be hashed, for example, and data-base updates handled piecemeal, with only the changes broadcast over a wide area. Reachability and delay metrics, if used, must not depend on direct connectivity to all other gateways or the use of network-specific broadcast mechanisms. Polling procedures (e.g. for consistency checking) should be used only sparingly and in no case introduce an overhead exceeding a constant independent of network topology times the longest non-looping path.

5. The use of a default gateway as a means to reduce the size of the routing data base is strongly discouraged in view of the many problems with multiple paths, loops and mis-configuration vulnerabilities. If used at all, it should be limited to a discovery function, with operational routes cached from external or internal data bases via either the routing algorithm or EGP.

6. This document places no restriction on the type of routing algorithm, such as node-based, link-based or any other algorithm, or metric, such as delay or hop-count. However, the size of the routing data base must not be allowed to exceed a constant independent of network topology times the

number of nodes times the mean connectivity (average number of incident links). An advanced design would not require that the entire routing data base be kept in any particular gateway, so that discovery and caching techniques would be necessary.

7. Operation and Maintenance

Gateways and packets switches are often operated as a system by some organization who agrees to operate and maintain the gateways, as well as to resolve link problems with the respective common carriers. It is important to note that the network control site may not be physically attached to the network being monitored. In general, the following requirements apply:

1. Each gateway must operate as a stand-alone device for the purposes of local hardware maintenance. Means must be available to run diagnostic programs at the gateway site using only on-site tools, which might be only a diskette or tape and local terminal. It is desirable, although not required, to run diagnostics via the network and to automatically reboot and dump the gateway via the net in case of fault. In general, this requires special hardware.

   The use of full-blown transport services such as TCP is in general ill-advised if required just to reboot and dump the gateway. Consideration should be given simple retransmission-overlay protocols based on UDP or specific monitoring protocols such as HMP described in RFC-869 [7].

2. It must be possible to reboot and dump the gateway manually from the control site. Every gateway must include a watchdog timer that either initiates a reboot or signals a remote control site if not reset periodically by the software. It is desirable that the data involved reside at the control site and be transmitted via the net; however, the use of local devices at the gateway site is acceptable. Nevertheless, the operation of initiating reboot or dump must be possible via the net, assuming a path is available and the connecting links are operating.

3. A mechanism must be provided to accumulate traffic statistics including, but not limited to, packet tallies, error-message tallies and so forth. The preferred method of retrieving these data is by explicit, periodic request from the control site using a standard datagram protocol based on UDP or HMP.

The use of full-blown transport services such as TCP is in general ill-advised if required just to collect statistics from the gateway. Consideration should be given simple retransmission-overlay protocols based on UDP or HMP.

4. Exception reports ("traps") occuring as the result of hardware or software malfunctions should be transmitted immediately (batched to reduce packet overheads when possible) to the control site using a standard datagram protocol based on UDP or HMP.

5. A mechanism must be provided to display link and node status on a continuous basis at the control site. While it is desirable that a complete map of all links and nodes be available, it is acceptable that only those components in use by the routing algorithm be displayed. This information is usually available locally at the control site, assuming that site is a participant in the routing algorithm.

The above functions require in general the participation of a control site or agent. The preferred way to provide this is as a user program suitable for operation in a standard software environment such as Unix. The program would use standard IP protocols such as TCP, UDP, and HMP to control and monitor the gateways. The use of specialized host hardware and software requiring significant additional investment is strongly discouraged; nevertheless, some vendors may elect to provide the control agent as an integrated part of the network in which the gateways are a part. If this is the case, it is required that a means be available to operate the control agent from a remote site using Internet protocols and paths and with equivalent functionality with respect to a local agent terminal.

Remote control of a gateway via Internet paths can involve either a direct approach, in which the gateway supports TCP and/or UDP directly, or an indirect approach, in which the control agent supports these protocols and controls the gateway itself using proprietary protocols. The former approach is preferred, although either approach is acceptable.

8. References and Bibliography

[1] Defense Advanced Research Projects Agency, "Internet Protocol", DARPA Network Working Group Report RFC-791, USC Information Sciences Institute, September 1981.

[2] Defense Advanced Research Projects Agency, "Internet Control Message Protocol", DARPA Network Working Group Report RFC-792, USC Information Sciences Institute, September 1981.

[3] Advanced Research Projects Agency, "Interface Message Processor - Specifications for the Interconnection of a Host and an IMP", BBN Report 1822, Bolt Beranek and Newman, December 1981.

[4] Plummer, D., "An Ethernet Address Resolution Protocol", DARPA Network Working Group Report RFC-826, Symbolics, September 1982.

[5] United States Department of Defense, "Military Standard Internet Protocol", Military Standard MIL-STD-1777, August 1983.

[6] Defense Communications Agency, "Defense Data Network X.25 Host Interface Specification", BBN Communications, December 1983.

[7] Hinden, R., "A Host Monitoring Protocol", DARPA Network Working Group Report RFC-869, BBN Communications, December 1983.

[8] Korb, J.T., "A Standard for the Transmission of IP Datagrams over Public Data Networks", DARPA Network Working Group Report RFC-877, Purdue University, September 1983.

[9] Nagle, J., "Congestion Control in IP/TCP Internetworks", DARPA Network Working Group Report RFC-896, Ford Aerospace, January 1984.

[10] Hornig, C., "A Standard for the Transmission of IP Datagrams over Ethernet Networks", DARPA Network Working Group Report RFC-894, Symbolics, April 1984.

[11] Mills, D.L., "Exterior Gateway formal Specification", DARPA Network Working Group Report RFC-904, M/A-COM Linkabit, April 1984.

[12] Postel, J., and J. Reynolds., "ARPA-Internet Protocol Policy", DARPA Network Working Group Report RFC-902, USC Information Sciences Institute, July 1984.

[13] Kirton, P., "EGP Gateway under Berkeley UNIX 4.2", DARPA Network Working Group Report RFC-911, USC Information Sciences Institute, August 1984.

[14] Postel, J., "Multi-LAN Address Resolution", DARPA Network Working Group Report RFC-925, USC Information Sciences Institute, October 1984.

[15] International Standards Organization, "Protocol for Providing the Connectionless-Mode Network Services", DARPA Network Working Group Report RFC-926, International Standards Organization, December 1984.

[16] National Research Council, "Transport Protocols for Department of Defense Data Networks", DARPA Network Working Group Report RFC-942, National Research Council, March 1985.

[17] Postel, J., "DOD Statement on NRC Report", DARPA Network Working Group Report RFC-945, USC Information Sciences Institute, April 1985.

[18] International Standards Organization, "Addendum to the Network Service Definition Covering Network Layer Addressing", DARPA Network Working Group Report RFC-941, International Standards Organization, April 1985.

[19] Leiner, B., J. Postel, R. Cole and D. Mills, "The DARPA Internet Protocol Suite", Proceedings INFOCOM 85, Washington DC, March 1985] Also in: IEEE Communications Magazine, March 1985.

[20] Winston, I., "Two Methods for the Transmission of IP Datagrams over IEEE 802.3 Networks", DARPA Network Working Group Report RFC-948, University of Pennsylvania, June 1985.

[21] Mogul, J., and J. Postel, "Internet Standard Subnetting Procedure", DARPA Network Working Group Report RFC-950, Stanford University, August 1985.

[22] Reynolds, J., and J. Postel, "Official ARPA-Internet Protocols", DARPA Network Working Group Report RFC-961, USC Information Sciences Institute, October 1985.

[23] Reynolds, J., and J. Postel, "Assigned Numbers", DARPA Network Working Group Report RFC-960, USC Information Sciences Institute, December 1985.

[24] Nagle, J., "On Packet Switches with Infinite Storage", DARPA Network Working Group Report RFC-970, Ford Aerospace, December 1985.

[25] Defense Communications Agency, "DDN Protocol Handbook", NIC-50004, NIC-50005, NIC-50006, (three volumes), SRI International, December 1985.

[26] Defense Communications Agency, "ARPANET Information Brochure", NIC-50003, SRI International, December 1985.

[27] Mills, D.L., "Autonomous Confederations", DARPA Network Working Group Report RFC-975, M/A-COM Linkabit, February 1986.

[28] Jacobsen, O., and J. Postel, "Protocol Document Order Information", DARPA Network Working Group Report RFC-980, SRI International, March 1986.

[29] Malis, A.G., "PSN End-to-End Functional Specification", DARPA Network Working Group Report RFC-979, BBN Communications, March 1986.

Appendix A. Ethernet Management

Following is a summary of procedures specified for use by hosts and gateways on an Ethernet.

A.1. Hardware

A packet is accepted from the cable only if its destination Ethernet address matches either the assigned interface address or a broadcast/multicast address. Presumably, this filtering is done by the interface hardware; however, the software driver is expected to do this if the hardware does not. Some hosts incorporate an optional feature that associates an assigned multicast address with a specific subnet in order to restrict access for testing, etc. When this feature is activated, the assigned multicast address replaces the broadcast address.

A.2. IP datagram

In case of broadcast/multicast (as determined from the destination Ethernet address) an IP datagram is discarded if the source IP address is not in the same subnet, as determined by the assigned host IP address and subnet mask. It is desirable that this test be overridden by a configuration parameter, in order to support the infrequent cases where more than one subnet may coexist on the same cable.

A.3. ARP datagram

An ARP reply is discarded if the destination IP address does not match the local host address. An ARP request is discarded if the source IP address is not in the same subnet. It is desirable that this test be overridden by a configuration parameter, in order to support the infrequent cases where more than one subnet may coexist on the same cable (see RFC-925 for examples). An ARP reply is generated only if the destination protocol IP address is reachable from the local host (as determined by the routing algorithm) and the next hop is not via the same interface. If the local host functions as a gateway, this may result in ARP replies for destinations not in the same subnet.

A.4. ICMP redirect

An ICMP redirect is discarded if the destination IP address does not match the local host address or the new target address is not on the same subnet. An accepted redirect updates the routing data base for the old target address. If there is no route or

associated with the old target address, the redirect is ignored. If the old route is associated with a default gateway, a new route associated with the new target address is inserted in the data base. Note that it is not possible to send a gratuitous redirect unless the sender is possessed of considerable imagination.

When subnets are in use there is some ambiguity as to the scope of a redirect, unless all hosts and gateways involved have prior knowledge of the subnet masks. It is recommended that the use of ICMP network-redirect messages be avoided in favor of ICMP host-redirect messages instead. This requires the original sender (i.e. redirect recipient) to support a general IP address-translation cache, rather than the usual network table. However, this is normally done anyway in the case of ARP.

An ICMP redirect is generated only if the destination IP address is reachable from the local host (as determined by the routing algorithm) and the next hop is via the same interface and the target address is defined in the routing data base. Redirects should never be sent in response to an IP net or subnet broadcast address or in response to a Class-D or Class-E IP address.

ICMP redirects are never forwarded, regardless of destination address. The source IP address of the ICMP redirect itself is not checked, since the sending gateway may use one of its addresses not on the common net. The source IP address of the encapsulated IP datagram is not checked on the assumption the host or gateway sending the original IP datagram knows what it is doing.

Appendix B. Policy Issues

The following sections discuss certain issues of special concern to the NSF scientific networking community. These issues have primary relevance in the policy area, but also have ramifications in the technical area.

B.1. Interconnection Technology

Currently the most important common interconnection technology between Internet systems of different vendors is Ethernet. Among the reasons for this are the following:

1. Ethernet specifications are well-understood and mature.

2. Ethernet technology is in almost all aspects vendor independent.

3. Ethernet-compatible systems are common and becoming more so.

These advantages combined favor the use of Ethernet technology as the common point of demarcation between NSF network systems supplied by different vendors, regardless of technology. It is a requirement of NSF gateways that, regardless of the possibly proprietary switching technology used to implement a given vendor-supplied network, its gateways must support an Ethernet attachment to gateways of other vendors.

It is expected that future NSF gateway requirements will specify other interconnection technologies. The most likely candidates are those based on X.25 or IEEE 802, but other technologies including broadband cable, fiber-optic or other protocols such as DDCMP may also be considered.

B.2. Proprietary and Extensible Issues

Internet technology is a growing, adaptable technology. Although hosts, gateways and networks supporting this technology have been in continuous operation for several years, vendors users and operators should understand that not all networking issues are fully understood. As a result, when new needs or better solutions are developed for use in the NSF networking community, it may be necessary to field new protocols. Normally, these new protocols will be designed to interoperate in all practical respects with existing protocols; however, occasionally it may happen that existing systems must be upgraded to support these protocols.

NSF systems vendors should understand that they also undertake a commitment to remain aware of current Internet technology and be prepared to upgrade their products from time to time as appropriate. As a result, these vendors are strongly urged to consider extensibility and periodic upgrades as fundamental characteristics of their products. One of the most productive and rewarding ways to do this on a long-term basis is to participate in ongoing Internet research and development programs in partnership with the academic community.

B.3. Multi-Protocol Gateways

Although the present requirements for an NSF gateway specify only the Internet protocol suite, it is highly desirable that gateway designs allow future extensions to support additional suites and allow simultaneous operation with more than a single one. Clearly, the ISO protocol suite is a prime candidate for one of these suites. Other candidates include XNS and DECnet.

Future requirements for NSF gateways may include provisions for other protocol suites in addition to Internet, as well as models and specifications to interwork between them, should that be appropriate. For instance, it is expected that the ISO suite will eventually become the dominant one; however, it is also expected that requirements to support other suites will continue, perhaps indefinitely.

Present NSF gateway requirements do not include protocols above the network layer, such as TCP, unless necessary for network monitoring or control. Vendors should recognize that future requirements to interwork between Internet and ISO applications, for example, may result in an opportunity to market gateways supporting multiple protocols at all levels through the application level. It is expected that the network-level NSF gateway requirements summarized in this document will be incorporated in the requirements document for these application-level gateways.

B.4. Access Control and Accounting

There are no requirements for NSF gateways at this time to incorporate specific access-control and accounting mechanisms in the design; however, these important issues are currently under study and will be incorporated into a redraft of this document at an early date. Vendors are encouraged to plan for the early introduction of these mechanisms in their products. While at this

time no definitive common model for access control and accounting has emerged, it is possible to outline some general features such a model is likely to have, among them the following:

1. The primary access control and accounting executive mechanisms will be in the service hosts themselves, not the gateways, packet switches or workstations.

2. Agents acting on behalf of access control and accounting executive mechanisms may be necessary in the gateways, packet switches or workstations. These may be used to collect data, enforce password protection or mitigate resource priority and fairness. However, the architecture and protocols used by these agents may be a local matter and not possible to specify in advance.

3. NSF gateways may be required to incorporate access control and accounting mechanisms based on packet source/destination address, as well as other fields in the IP header, internal priority and fairness. However, it is extremely unlikely that these mechanisms would involve a user-level login to the gateway itself.